Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

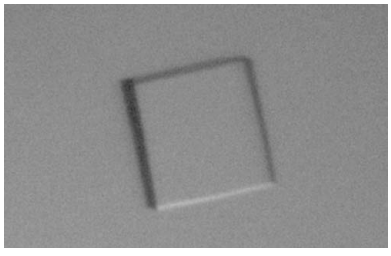
The production and crystallization of the human leukocyte antigen class II molecules HLA-DQ2 and HLA-DQ8 complexed with deamidated gliadin peptides implicated in coeliac disease

Kate N. Henderson,[a] Hugh H. Reid,[a] Natalie A. Borg,[a] Sophie E. Broughton,[a] Trevor Huyton,[a] Robert P. Anderson,[b,c] James McCluskey[d] and Jamie Rossjohn[a]*

[a]The Protein Crystallography Unit, Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia, [b]Autoimmunity and Transplantation Division, Walter and Eliza Hall Institute, 1G Royal Parade, Parkville, Victoria 3050, Australia, [c]Department of Gastroenterology, The Royal Melbourne Hospital, Grattan Street, Parkville, Victoria 3050, Australia, and [d]Department of Microbiology and Immunology, University of Melbourne, Parkville, Victoria 3010, Australia

Correspondence e-mail: jamie.rossjohn@med.monash.edu.au

Received 2 August 2007
Accepted 17 October 2007



The major histocompatibility complex (MHC) class II molecules HLA-DQ2 and HLA-DQ8 are key risk factors in coeliac disease, as they bind deamidated gluten peptides that are subsequently recognized by $CD4^+$ T cells. Here, the production and crystallization of both HLA-DQ2 and HLA-DQ8 in complex with the deamidated gliadin peptides DQ2 α-I (PQPELPYPQ) and DQ8 α-I (EGSFQPSQE), respectively, are reported.

1. Introduction

The MHC class II glycoproteins are heterodimers consisting of an α and a β chain that associate together noncovalently to form a peptide-binding groove. This groove is lined with pockets that are able to accommodate peptide side chains. The residues lining these pockets are extremely polymorphic and are responsible for conferring the different peptide-binding preferences exhibited by different alleles. In contrast to other HLA (human leukocyte antigen) molecules, the peptide-binding motif for both HLA-DQ2 and HLA-DQ8 is characterized by a preference for negatively charged residues at anchor positions 4, 6 and 7 for HLA-DQ2 and at positions 1 and 9 for HLA-DQ8 (Kim *et al.*, 2004; van de Wal *et al.*, 1996; Vartdal *et al.*, 1996; Godkin *et al.*, 1997; Kwok *et al.*, 1996; Lee *et al.*, 2001).

Coeliac disease (CD) is a gluten-induced autoimmune-like disease of the small intestine that affects caucasians, with an incidence rate approaching 1% (Fasano *et al.*, 2003; Maki *et al.*, 2003; Catassi *et al.*, 1994; Csizmadia *et al.*, 1999; Rostami *et al.*, 1999). The MHC molecules HLA-DQ2 and HLA-DQ8 are key risk factors in the disease, with 90% of coeliac patients being $HLA\text{-}DQ2^+$ whilst the remainder are generally $HLA\text{-}DQ8^+$ (Marsh, 1992; Tighe *et al.*, 1992; Sollid *et al.*, 1989; Spurkland *et al.*, 1997). In CD, HLA-DQ2 and HLA-DQ8, which are present on antigen-presenting cells, contribute to disease pathogenesis by binding gluten peptides, which are then recognized by $CD4^+$ T cells.

Since gluten, which has a high proline and glutamine content, contains very few negative residues, it was unclear how gluten peptides could be bound by HLA-DQ2 and HLA-DQ8. Subsequently, it was discovered that the gluten-specific $CD4^+$ T cells predominantly recognize modified gluten peptides in which selected glutamines (Q) have been converted to negatively charged glutamates (E) (Molberg *et al.*, 2001; van de Wal *et al.*, 1998). This deamidation reaction, catalysed by the enzyme tissue transglutaminase (TG2), generally increases the immunogenicity of the gluten peptides, presumably by enabling the gluten peptides to be bound by HLA-DQ2 and HLA-DQ8 (van de Wal *et al.*, 1998).

In order to determine the structural basis for the increased immunogenicity of deamidated peptides in CD, we have crystallized HLA-DQ2 and HLA-DQ8 with the deamidated gliadin peptides DQ2 α-II (PQPELPYPQ) and DQ8 α-I (EGSFQPSQE), respectively, from the α-gliadin family. Presently, there is only one report (Kim *et al.*, 2004) that has investigated the binding of a deamidated peptide to HLA-DQ2; no structural studies have investigated the binding of these peptides to HLA-DQ8. Accordingly, more studies

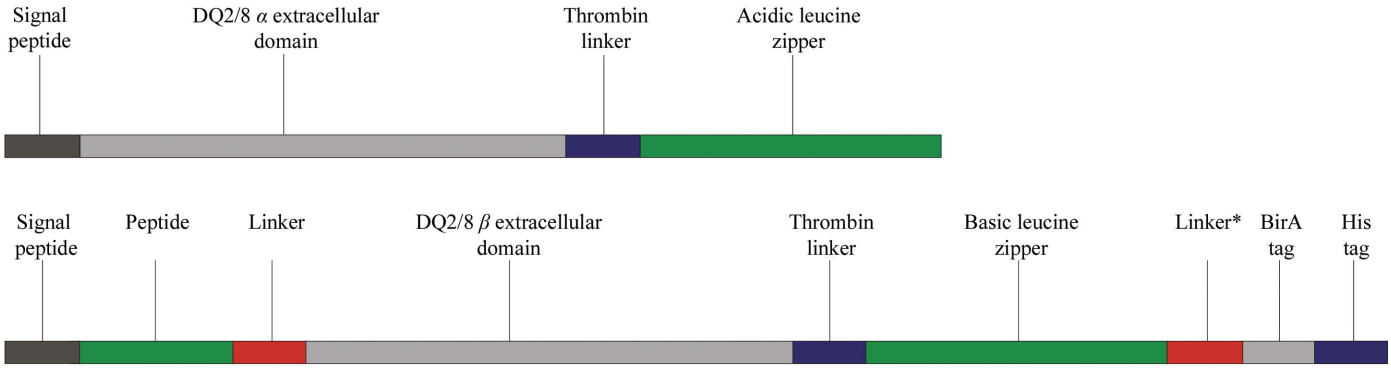


Figure 1
The design of the HLA-DQ α-chain construct (*a*) and β-chain construct (*b*). The constructs were made by overlapping PCR and cloned into the dual expression vector pFastBac Dual for simultaneous expression. Note that the HLA-DQ2 β-chain construct contained an additional flexible linker that is marked by an asterisk.

are required to fully appreciate the deamidation-dependence of gluten peptides in CD.

2. Methods

2.1. Construct design

The design of the HLA-DQ2/8 expression constructs (Fig. 1) was based on previously developed methods for the production of soluble MHC class II molecules (Kozono *et al.*, 1994; Scott *et al.*, 1996; Crawford *et al.*, 1998). Constructs were designed to express the mature extracellular domains of the α (HLA-DQ2 residues 1–193, HLA-DQ8 residues 1–184) and β (HLA-DQ2 residues 1–198, HLA-DQ8 residues 1–192) chains attached *via* thrombin cleavage linkers (SSADLVPRGS) to a pair of leucine zippers (acidic zipper sequence TTAPSAQLEKELQALQKENAQLEWELQALEKELAQ, basic zipper sequence TTAPSAQLKKKLQALKKKNAQLKWKLQALKKKLAQ). The leucine zippers facilitated α- and β-chain dimerization and also aided in protein stabilization. In addition, the β-chain construct also contained a BirA biotinylation tag (LGGIFEAMKMELRD) for tetramer production, necessary only for producing MHC class II tetramers which were used in other studies not described here, and a histidine tag (ten-residue tag for HLA-DQ2, six-residue tag for HLA-DQ8) for purification purposes. The gliadin peptides (HLA-DQ2 peptide sequence QLQKFPQPELPYPQPQP, HLA-DQ8 peptide sequence QQYPSGEGSFQPSQENPQ) were attached *via* a flexible glycine-serine linker containing a factor X cleavage site (GGGGSIEGRGSGGGS) to the N-terminus of the β chain, ensuring proper loading of the peptide into the MHC binding groove. Further modifications of the HLA-DQ2 construct included mutation of a free cysteine to serine (α47 C$\rightarrow$S) to increase protein stability and the addition of an extra glycine-serine linker (GSGSGSGS) to increase tetramer flexibility. Full sequence details are provided as supplementary material[1].

2.2. PCR and cloning

The HLA-DQ2/8 constructs were produced using overlapping PCR. A full description of the primers and the overlapping PCR reactions performed are detailed in the supplementary material[1]. HLA-DQ2 cDNA was a gift from Bill Kwok and Gerald Nepom, HLA-DQ8 cDNA was a gift from Zhenjun Chen and leucine-zipper cDNA was a gift from Takehiko Sasazuki and Yoshinori Fukui. Primers were purchased from Geneworks and various DNA polymerases were used.

2.3. Generation of recombinant viruses

The Bac-to-Bac Baculovirus Expression System (Invitrogen) was used to express the HLA-DQ2/8 α- and β-chain constructs in High Five insect cells according to the manufacturer's instructions. The HLA-DQ α- and β-chain constructs were cloned into the expression vector pFastBac Dual and recombinant baculoviruses were generated using Sf9 insect cells.

2.4. Protein expression and purification

HLA-DQ2$^{\mathrm{PQPELPYPQ}}$ and HLA-DQ8$^{\mathrm{EGSFQPSQE}}$ were expressed in High Five cells as a soluble protein secreted into the medium. High Five cells were grown at 300 K in shaker flasks using serum-free medium (Hyclone, HyQ SFX-Insect). For recombinant protein expression, High Five cells at 2.5×10^6 cells ml^{-1} were infected with recombinant virus with a multiplicity of infection of five and the medium was harvested 70 h later by centrifugation at 1500*g* (Sorvall Evolution RC centrifuge). The medium was concentrated and buffer-exchanged into 10 m*M* Tris–HCl pH 8.0, 10 m*M* imidazole, 500 m*M* NaCl using a Sarton Slice Crossflow System (Sartorious) with a Hydrosart cassette (Sartorious, 10 kDa cutoff). The DQ2/8 protein constructs were captured onto Ni–NTA metal-affinity resin (Qiagen) and eluted using 10 m*M* Tris–HCl pH 8.0, 300 m*M* imidazole. The proteins were further purified by anion exchange with a 5 ml Q Sepharose High Performance HiTrap column (Amersham Biosciences) using a salt gradient. The acidic and basic zippers were removed by digestion using trypsin-agarose beads (Sigma) and the zippers were purified away by size exclusion using a HiLoad Superdex 75 16/60 column (Amersham Biosciences) in 10 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl. The final products for crystallization are defined precisely in the supplementary information[1] in terms of DNA sequences. Proteins were concentrated and quantified by spectrophotometric analysis. The protein purity was monitored using SDS–PAGE. The protease inhibitor phenylmethylsulfonyl fluoride (PMSF) was used throughout the purification process.

2.5. Crystallization and data collection

Initial crystallization trials were performed at room temperature by the sitting-drop vapour-diffusion method using nanodrops (50 nl protein solution and 50 nl mother liquor) dispensed by a nanopipetting robot (Cartesian Technologies) into a 96-well Intelli-Plate

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: EN5259).

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

	HLA-DQ2^{PQPELPYPQ}	HLA-DQ8^{EGSFQPSQE}
X-ray source	GM/CA-CAT APS	Rigaku RU-H3RHB
Detector	MAR 300	R-AXIS IV++
Temperature (K)	100	100
Space group	$I23$	$P2_12_12_1$
Unit-cell parameters (Å)	a = 135.86	a = 58.83, b = 92.52, c = 113.06
Resolution (Å)	48.11–3.90 (4.11–3.90)	49.64–2.10 (2.18–2.10)
Total No. of observations	47405 (6959)	126773 (12379)
No. of unique observations	3946 (565)	69568 (6966)
Multiplicity	12.0 (12.3)	1.84 (1.78)
Data completeness (%)	100 (100)	99.2 (99.9)
$I/\sigma(I)$	5.5 (1.8)	8.6 (2.1)
R_{merge}† (%)	8.6 (40.9)	5.6 (27.1)
Matthews coefficient (V_M) (Å^3 Da^{-1})	2.06	3.02
Solvent content (%)	40	59
Wilson B factor (Å^2)	100	37.9

† R_{merge} measures the agreement among multiple measurements of the same reflections and is calculated as follows: $R_{merge} = \sum |I_{hkl} - \langle I_{hkl}\rangle| / \sum I_{hkl}$.

(Art Robbins Instruments). Initial crystallization conditions were established using commercial screens, with protein concentrations of 4.5 mg ml^{-1} for HLA-DQ2^{PQPELPYPQ} and 7 mg ml^{-1} for HLA-DQ8^{EGSFQPSQE}, both in 10 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl. Both proteins were screened in Crystal Screen 1 and PEG/Ion Screen 1 (Hampton Research), while HLA-DQ2^{PQPELPYPQ} was additionally screened in Index Screen (Hampton Research) and JB Screen Classic 1–4 (Jena Bioscience). Fine screens were set up around the most promising conditions using the hanging-drop vapour-diffusion method at room temperature in 24-well Linbro tissue-culture plates.

The final crystallization conditions for HLA-DQ2^{PQPELPYPQ} were obtained with 1 µl droplets prepared by mixing equal volumes of protein (4.5 mg ml^{-1} in 10 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl) and reservoir solution [0.1 *M* bis-Tris pH 6.2, 20–25%(*w*/*v*) polyethylene glycol (PEG) 3350]. A reservoir volume of 1 ml was used. Cubic crystals grew within 1–2 weeks. Crystals were dehydrated in a capillary containing the precipitant buffer with a higher concentration of precipitant [30%(*w*/*v*) PEG 3350] for at least 3 d before being flash-cooled in liquid nitrogen.

The final crystallization conditions for HLA-DQ8^{EGSFQPSQE} were obtained with 2 µl droplets prepared by mixing equal volumes of protein (7 mg ml^{-1} in 10 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl) and reservoir solution [0.05 *M* monopotassium dihydrogen phosphate, 19–21%(*w*/*v*) PEG 8000]. A reservoir volume of 1 ml was used. Crystals with rod-like morphology grew within a week. The crystals were flash-cooled in nitrogen gas prior to data collection using 20%(*v*/*v*) glycerol in the precipitant buffer as the cryoprotectant.

A complete data set for HLA-DQ2^{PQPELPYPQ} was collected at the GM/CA-CAT ID-D beamline using a MAR 300 detector (Advanced Photon Source, Chicago). A total of 200 images were collected (0.5° oscillation, 1 s exposure per oscillation) and the data were processed and scaled using the programs *MOSFLM* v.6.2.6 (Leslie, 1992) and *SCALA* (Kabsch, 1988) from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). A complete data set for HLA-DQ8^{EGSFQPSQE} was collected in-house on an R-AXIS IV^{++} detector with Cu $K\alpha$ X-rays generated by a Rigaku RU-H3RHB rotating-anode generator and focused using Osmic mirrors. A total of 173 images were collected (0.5° oscillation, 15 min exposure per oscillation) and the data were processed with *d*TREK* (Pflugrath, 1999).

3. Results and discussion

3.1. Cloning, expression and purification

The engineered HLA-DQ2^{PQPELPYPQ} and HLA-DQ8^{EGSFQPSQE} constructs were produced by overlapping PCR and cloned into pFastBac Dual for expression in High Five cells. Protein was purified from the medium by metal-affinity and anion-exchange chromatography. Anion-exchange purification of HLA-DQ8^{EGSFQPSQE} resulted in one peak being eluted, whilst multiple peaks were eluted for HLA-DQ2^{PQPELPYPQ}. Interestingly, it was found that once the leucine zippers had been removed, HLA-DQ2^{PQPELPYPQ} also eluted as one peak. Leucine zippers were removed by trypsin digestion followed by size-exclusion chromatography. Although the HLA-DQ2/8 constructs were engineered with a thrombin recognition site for this purpose, trypsin was found to cleave this site much more effectively, resulting in 100% cleavage as evidenced by the complete shift of both the α and β chains on a reducing SDS–PAGE gel (Fig. 2). Pre- and post-cleavage the HLA-DQ2/8 protein runs as a doublet, with the molecular weight of the β chains corresponding to the theoretical molecular weight (pre-cleavage β chain $\simeq$ 35 kDa, post-

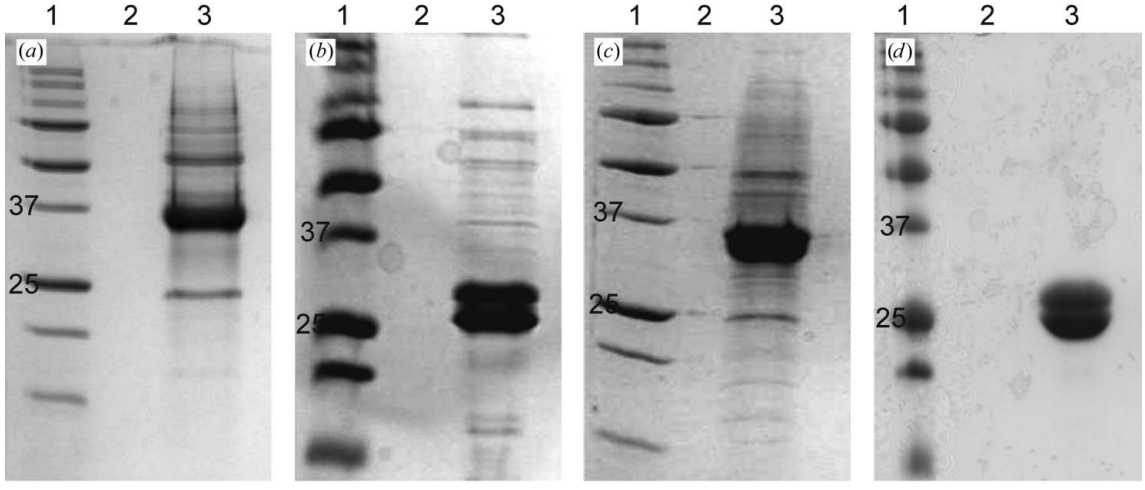


Figure 2
Reducing SDS–PAGE analysis of purified HLA-DQ2^{PQPELPYPQ} (*a*) prior to trypsin cleavage and (*b*) post trypsin cleavage and of purified HLA-DQ8^{EGSFQPSQE} (*c*) prior to trypsin cleavage and (*d*) post trypsin cleavage. MW markers (kDa) are in lane 1 and purified protein is in lane 3.

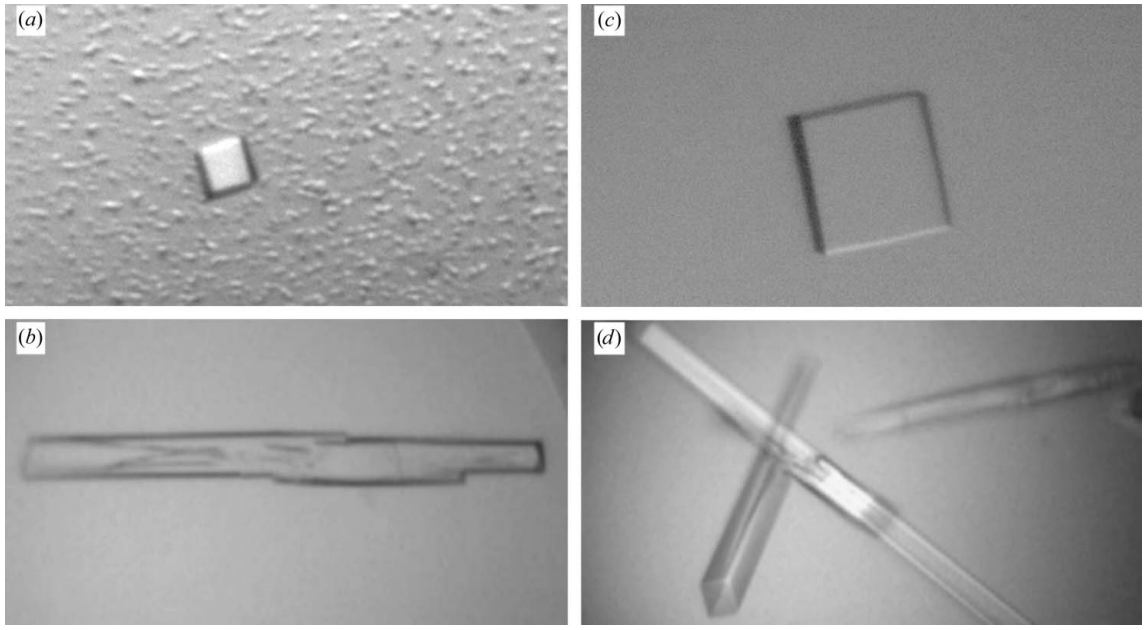


Figure 3
(*a*) Preliminary and (*b*) diffraction-quality crystals of HLA-DQ2^{PQPELPYPQ}; (*c*) preliminary and (*d*) diffraction quality crystals of HLA-DQ8^{EGSFQPSQE}.

cleavage $\simeq$ 28 kDa) whilst the α chain ran higher than expected (pre-cleavage α chain $\simeq$ 27 kDa, post cleavage $\simeq$ 23 kDa) (Fig. 2). As seen in Fig. 2, SDS–PAGE analysis of the α and β chains resulted in the presence of smeary bands or at times the presence of multiple bands, which is mostly likely to be a consequence of glycosylation of the two chains. N-terminal sequencing and Western blot analysis with an antihistidine antibody confirmed the correct identity of the proteins. The protein was concentrated prior to crystallization trials, with the final protein yields being 0.05 and 0.1 mg per litre of culture for HLA-DQ2^{PQPELPYPQ} and HLA-DQ8^{EGSFQPSQE}, respectively. Owing to the low protein yield, further purification of HLA-DQ2^{PQPELPYPQ} was not pursued. HLA-DQ2^{PQPELPYPQ} purity was ~90%, whilst HLA-DQ8^{EGSFQPSQE} purity was >95%, as estimated by SDS–PAGE (Fig. 2).

3.2. Crystallization and data collection

Initial crystallization trials using commercial screens identified several conditions in which the HLA-DQ2^{PQPELPYPQ} protein crystallized. Condition 43 of Index Screen 1 [0.1 *M* bis-tris pH 6.5, 25%(*w*/*v*) PEG 3350] produced small but nicely formed discrete crystals within 3 d (Fig. 3). Crystals were able to be reproduced by performing a fine PEG/pH screen around the condition; on lowering the precipitant concentration, larger crystals grew over a few weeks. Initially, crystals were flash-cooled using 5%(*v*/*v*) glycerol in the precipitant buffer as the cryoprotectant; however, these crystals were found to diffract poorly at ~9 Å. By dehydrating the crystals in precipitant buffer containing a higher concentration of PEG [30%(*w*/*v*)] prior to data collection, the diffraction limit was markedly improved. A 3.9 Å data set was collected from a crystal belonging to the cubic space group *I*23, with unit-cell parameter 135.86 Å. The Matthews coefficient (V_M) of 2.06 Å^3 Da^{-1} and a solvent content of 40% were consistent with the presence of one molecule per asymmetric unit (Matthews, 1968; Collaborative Computational Project, Number 4, 1994). Unfortunately, further dehydration experiments failed to improve the diffraction quality of these crystals. Like HLA-DQ2^{PQPELPYPQ}, HLA-DQ8^{EGSFQPSQE} also crystallized in a number of conditions using the commercial screens. Most of these conditions resulted in needle-like crystals; however, condition 42 of Crystal Screen 1 [20%(*w*/*v*) PEG 8000, 0.05 *M* potassium phosphate] produced a discrete but ill-formed crystal (Fig. 3). Larger, better formed crystals with rod-shaped morphology were produced reproducibly by performing a fine PEG screen around this condition (Fig. 3). A single crystal diffracted to 2.1 Å and belonged to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 58.83, b = 92.52, c = 133.06 Å. The Matthews coefficient (V_M) of 3.02 Å^3 Da^{-1} and a solvent content of 59% were consistent with the presence of one molecule per asymmetric unit (Matthews, 1968; Collaborative Computational Project, Number 4, 1994). Data-collection statistics for both HLA-DQ2^{PQPELPYPQ} and HLA-DQ8^{EGSFQPSQE} are shown in Table 1.

In this report, the production and crystallization of HLA-DQ2^{PQPELPYPQ} and HLA-DQ8^{EGSFQPSQE} in complex with deamidated gliadin peptides has been described. Further crystallization trials are under way for HLA-DQ2^{PQPELPYPQ}, whereas for HLA-DQ8^{EGSFQPSQE} structure determination has been completed (PDB code 2nna). The structures of HLA-DQ2 and HLA-DQ8 in complex with deamidated gliadin peptides will be vital to understanding the role of deamidation in CD and should facilitate the design of novel ways to treat and manage the disease. Additionally, valuable insights will also be gained into post-translational modification of antigens involved in other autoimmune diseases.

This work was supported by the NHMRC and by an ARC Federation Fellowship to JR and a Monash University Australian Postgraduate Award to KH.

References

Catassi, C., Ratsch, I. M., Fabiani, E., Rossini, M., Bordicchia, F., Candela, F., Coppa, G. V. & Giorgi, P. L. (1994). *Lancet*, **343**, 200–203.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Crawford, F., Kozono, H., White, J., Marrack, P. & Kappler, J. (1998). *Immunity*, **8**, 675–682.

Csizmadia, C. G., Mearin, M. L., von Blomberg, B. M., Brand, R. & Verloove-Vanhorick, S. P. (1999). *Lancet*, **353**, 813–814.
Fasano, A. *et al.* (2003). *Arch. Intern. Med.* **163**, 286–292.
Godkin, A., Friede, T., Davenport, M., Stevanovic, S., Willis, A., Jewell, D., Hill, A. & Rammensee, H. G. (1997). *Int. Immunol.* **9**, 905–911.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.
Kim, C. Y., Quarsten, H., Bergseng, E., Khosla, C. & Sollid, L. M. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 4175–4179.
Kozono, H., White, J., Clements, J., Marrack, P. & Kappler, J. (1994). *Nature (London)*, **369**, 151–154.
Kwok, W. W., Domeier, M. L., Raymond, F. C., Byers, P. & Nepom, G. T. (1996). *J. Immunol.* **156**, 2171–2177.
Lee, K. H., Wucherpfennig, K. W. & Wiley, D. C. (2001). *Nature Immunol.* **2**, 501–507.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Maki, M., Mustalahti, K., Kokkonen, J., Kulmala, P., Haapalahti, M., Karttunen, T., Ilonen, J., Laurila, K., Dahlbom, I., Hansson, T., Hopfl, P. & Knip, M. (2003). *N. Engl. J. Med.* **348**, 2517–2524.
Marsh, M. N. (1992). *Gastroenterology*, **102**, 330–354.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–487.
Molberg, O., McAdam, S., Lundin, K. E., Kristiansen, C., Arentz-Hansen, H., Kett, K. & Sollid, L. M. (2001). *Eur. J. Immunol.* **31**, 1317–1323.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Rostami, K., Mulder, C. J., Werre, J. M., van Beukelen, F. R., Kerchhaert, J., Crusius, J. B., Pena, A. S., Willekens, F. L. & Meijer, J. W. (1999). *Scand. J. Gastroenterol.* **34**, 276–279.
Scott, C. A., Garcia, K. C., Carbone, F. R., Wilson, I. A. & Teyton, L. (1996). *J. Exp. Med.* **183**, 2087–2095.
Sollid, L. M., Markussen, G., Ek, J., Gjerde, H., Vartdal, F. & Thorsby, E. (1989). *J. Exp. Med.* **169**, 345–350.
Spurkland, A., Ingvarsson, G., Falk, E. S., Knutsen, I., Sollid, L. M. & Thorsby, E. (1997). *Tissue Antigens*, **49**, 29–34.
Tighe, M. R., Hall, M. A., Barbado, M., Cardi, E., Welsh, K. I. & Ciclitira, P. J. (1992). *Tissue Antigens*, **40**, 90–97.
Vartdal, F., Johansen, B. H., Friede, T., Thorpe, C. J., Stevanovic, S., Eriksen, J. E., Sletten, K., Thorsby, E., Rammensee, H. G. & Sollid, L. M. (1996). *Eur. J. Immunol.* **26**, 2764–2772.
Wal, Y. van de, Kooy, Y., Drijfhout, J. W., Amons, R. & Koning, F. (1996). *Immunogenetics*, **44**, 246–253.
Wal, Y. van de, Kooy, Y., van Veelen, P., Pena, S., Mearin, L., Papadopoulos, G. & Koning, F. (1998). *J. Immunol.* **161**, 1585–1588.